Predicting Coronal Mass Ejection Arrival Times with Thirty-Minute Accuracy

Gabor Toth[1]

[1]University of Michigan-Ann Arbor

January 17, 2023

Abstract

J. Schmidt and Cairns (2019) have recently shown that they can predict Coronal Mass Ejection (CME) arrival times with an accuracy of 0.9+-1.9 hours for four separate events. They also showed that the accuracy gets better with increased grid resolution. Here, we further improve these results by using the Richardson extrapolation (Richardson and Gaunt, 1927), which is a standard technique in computational fluid dynamics, and predict the CME arrival time with 0.2+-0.26 hours accuracy. The CME arrival time errors of the new model lie in a 95% confidence interval [-0.21,0.61] h. We also show that the probability of getting these accurate arrival time predictions with a model with a standard deviation exceeding 2 hours is less than 0.1%, indicating that the excellent results cannot be due to random chance, and the Richardson extrapolation has indeed improved the original model by J. Schmidt and Cairns (2019). This unprecedented accuracy is about 40 times better than the current state-of-the-art prediction of CME arrival times with an average error of about +-10 hours. The new model uses information available within a few hours after the CME eruption and it can run much faster than real-time on a couple of CPU cores. Based on the result, we recommend the new model to be transitioned to operations as soon as possible to better protect our space-born and ground-based assets from the harmful effects of space weather.

Predicting Coronal Mass Ejection Arrival Times with Thirty-Minute Accuracy

Gábor Tóth

Dept. of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA

Abstract

J. Schmidt and Cairns (2019) have recently shown that they can predict Coronal Mass Ejection (CME) arrival times with an accuracy of 0.9±1.9 hours for four separate events. They also showed that the accuracy gets better with increased grid resolution. Here, we further improve these results by using the Richardson extrapolation (Richardson & Gaunt, 1927), which is a standard technique in computational fluid dynamics, and predict the CME arrival time with 0.2 ± 0.26 hours accuracy. The CME arrival time errors of the new model lie in a 95% confidence interval $[-0.21, 0.61]$ h. We also show that the probability of getting these accurate arrival time predictions with a model with a standard deviation exceeding 2 hours is less than 0.1%, indicating that the excellent results cannot be due to random chance, and the Richardson extrapolation has indeed improved the original model by J. Schmidt and Cairns (2019). This unprecedented accuracy is about 40 times better than the current state-of-the-art prediction of CME arrival times with an average error of about ±10 hours. The new model uses information available within a few hours after the CME eruption and it can run much faster than real-time on a couple of CPU cores. Based on the result, we recommend the new model to be transitioned to operations as soon as possible to better protect our space-born and ground-based assets from the harmful effects of space weather.

1 Introduction

Predicting the propagation of Coronal Mass Ejections (CMEs) and their arrival time at Earth has been a major goal of space weather prediction for decades. The ENLIL model (Odstrčil & Pizzo, 1999a, 1999b), for example, solves the ideal magnetohydrodynamic (MHD) equations from about $20\,R_s$ (solar radii) to the Earth orbit and beyond. The inner boundary conditions are provided by the Wang-Sheeley-Arge (WSA) model (Arge & Pizzo, 2000). CMEs are initiated with the empirical CONE model based on flare observations and coronal white light images. Another approach is followed by the Alfvén Wave Solar atmosphere Model (AWSoM) (van der Holst et al., 2014) that is based on the BATS-R-US MHD code (Powell et al., 1999; Tóth et al., 2012), also widely used to model the solar corona, the heliosphere and the eruption and propagation of CMEs from the surface of the Sun (initiated by a flux rope model) to Earth and beyond. AWSoM solves the MHD equations extended with solar wind heating and acceleration due to Alfvén wave turbulence, radiative cooling and heat conduction. However, these first-principles models can only achieve about 10-hour accuracy predicting the CME arrival time (Wold et al., 2018, cf.). More recently, empirical and neural network based models were applied to this problem, but the typical error remains about ± 10 hours (Riley et al., 2018; Amerstorfer et al., 2021, cf.).

J. Schmidt and Cairns (2019) have used an earlier version of BATS-R-US developed by Cohen et al. (2007), which relies on a spatially varying polytropic index derived from the Wang-Sheely-Arge (WSA) model (Arge & Pizzo, 2000) and achieved an unprecedented accuracy for predicting the CME arrival time: 0.9 ± 1.9 hours. They describe their procedure of setting up the CME simulations using only information that is available prior to and within a few hours after the CME eruptions: the WSO magnetogram, the CME speed estimated from the CME Analysis Tool (CAT) using STEREO/LASCO C3 coronagraph images, and prior L1 observations used for the WSA model and in turn for BATS-R-US. In addition, we have learned from the authors that the simulations were performed on a couple of CPU cores and they managed to run the model about three times faster than real time. This means that their procedure is undoubtedly suitable for operational use.

Unfortunately, J. Schmidt and Cairns (2019) have only published their work in form of a preprint on arxiv, so it did not get the attention that it deserves. In fact, their work has been peer-reviewed and accepted for publication in one of the leading journals, and the actual publication was only delayed for some minor technical reasons. As it is explained by J. Schmidt and Cairns (2019), the "setup and analysis is refined from our earlier work simulating type II radio bursts and CMEs", which in fact resulted in four peer-reviewed and published works (J. M. Schmidt et al., 2013; J. M. Schmidt & Cairns, 2014, 2016; J. M. Schmidt et al., 2016). Therefore the research contained in (J. Schmidt & Cairns, 2019) can be safely considered to have the same quality and rigor as these prior publications. It is therefore imperative to make sure that this research is not forgotten and benefit the entire space weather community.

Looking at Figure 4 in (J. Schmidt & Cairns, 2019), reproduced here as Figure 1, we have noticed that the distances between the observations (diamonds) and the model predictions obtained on two different computational grids (squares and stars) form a distinctive pattern: the distances between the three symbols appear to be approximately

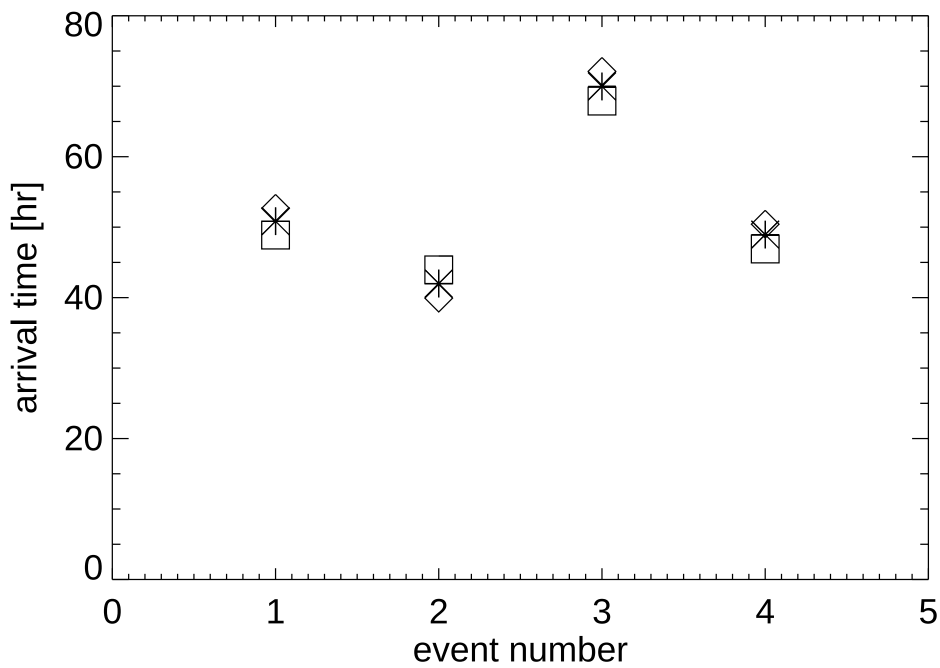

Figure 1. Observed and predicted arrival times at 1 au of four CME events (4 Sep 2017, 6 Sep 2017, 12 Feb 2018, and 29 Nov 2013 CME) recreated from (J. Schmidt & Cairns, 2019). The diamonds show observed arrival times, the squares and stars are simulation results at level 2 and level 5 grid refinements, respectively.

the same for all four events displayed. The main idea in this paper is that this fact can be exploited to obtain an even more accurate estimate of the CME arrival time.

The new arrival time estimate based on the Richardson extrapolation (Richardson & Gaunt, 1927) has a bias and standard deviation of 0.2 ± 0.26 hours, which is significantly better than the 0.9 ± 1.9 hours obtained by J. Schmidt and Cairns (2019). We will also show that the agreement between observations and simulations cannot be attributed to luck. Since the four events happened in different years and/or have very different arrival times covering a wide range from about 40 hours to 72 hours, the technique must be applicable to most CMEs. This means that the model should provide extremely reliable and accurate information for operational space weather forecasters, which is important for our national security and human safety.

2 Predicting CME arrival times

To perform a quantitative evaluation of the results presented in Figure 4 of (J. Schmidt & Cairns, 2019), we have digitized the figure and put the observed and simulated arrival times (relative to the eruption time) into Table 1. These values were also used to produce Figure 1 confirming that the values were extracted correctly.

Table 1. Simulated and observed CME arrival times for four events from Figure 4 in (J. Schmidt & Cairns, 2019). The times are measured in hours from the eruption time. The error is the absolute value of the difference between the observed and simulated times.

ID	Date	Observed	Model1	Model2	Error1	Error2	Error1/Error2
1	Sep 04, 2017	52.676	48.873	50.845	3.803	1.831	2.077
2	Sep 06, 2017	39.929	43.943	42.000	4.014	2.071	1.938
3	Feb 12, 2018	72.112	67.887	69.979	4.225	2.133	1.981
4	Nov 29, 2013	50.422	46.901	48.943	3.521	1.479	2.381
Average					3.891	1.8735	2.094

The errors Error1 and Error2 of the two models Model1 and Model2, corresponding to Refinement Level 2 and 5 in (J. Schmidt & Cairns, 2019), are remarkably constant across the four events, and the ratio of the errors is approximately 2.1. Using the idea of the Richardson extrapolation, which improves the numerical accuracy by estimating the exact solution from numerical solutions at two different grid resolutions, we construct the following formula for the extrapolated arrival time:

$$T_R = 2T_2 - T_1 \tag{1}$$

where T_1 and T_2 are the arrival times predicted by models 1 and 2. This provides the smallest error compared to observations and it is also consistent with the Richardson extrapolation for a first order accurate scheme. Indeed, the numerical scheme for discontinuities, like the CME shock, is only first order accurate.

3 Statistical Analysis and Probability Estimates

Table 2 shows that the mean absolute error of the extrapolated arrival time is about 0.218 hours, which is useful information, but not suitable for statistical analysis. To bet-

Table 2. Observed and extrapolated CME arrival times for four events. The times are measured in hours from the eruption time. The last column is the absolute value of the error.

ID i	Date	Observed T_i	Extrapolated $T_{i,R}$	Error $T_{i,R} - T_i$
1	Sep 04, 2017	52.676	52.817	0.141
2	Sep 06, 2017	39.929	40.057	0.128
3	Feb 12, 2018	72.112	72.071	-0.041
4	Nov 29, 2013	50.422	50.985	0.562
Mean absolute error				0.218
Mean $\pm$ one standard deviation				0.198 ± 0.257

ter quantify the performance of the new model, we calculate an unbiased estimate and a 95% confidence interval for the arrival time errors.

The sample size is $N = 4$. The average of the errors, the bias, is

$$B = \frac{1}{N}\sum_{i=1}^{N}(T_{i,R} - T_i) = 0.198\,h \tag{2}$$

and the sample variance S is

$$S = \sqrt{\frac{\sum_{i=1}^{N}(T_{i,R} - T_i - B)^2}{N-1}} = 0.257\,\mathrm{h} \tag{3}$$

where T_i is the observed arrival time for event i and $T_{i,R}$ is the Richardson extrapolated time calculated from Equation 1. The 95% confidence interval for the error $T_R - T$ is $B \pm tS/\sqrt{N}$, where $t = 3.182$ from the T-distribution for $p = 0.025$ and $N - 1 = 3$ degrees of freedom:

$$(T_R - T) \in [-0.211, 0.607]\,\mathrm{h} \tag{4}$$

We conclude that there is a 95% chance that the model will produce arrival time predictions with errors less than 37 minutes, while the average error is only 12 minutes.

Finally, it is important to check if the small errors in Table 2 are statistically significant, or they can be attributed to simple luck. Let us assume that the new model with the extrapolation has no bias, $\mu = 0$, and its standard deviation is $\sigma = 2\,\mathrm{h}$. The quantity

$$X^2 = \frac{\sum_{i=1}^{N}(T_{i,R} - T_i)^2}{\sigma^2} = 0.0887 \tag{5}$$

follows the $\chi^2(N, p)$ distribution since the mean value is assumed to be known. For $N = 4$, we find that there is only $p = 0.1\%$ chance that $X^2 \leq 0.0887$ by pure luck. If σ was larger than 2 hours, this probability would be even less. We can safely conclude that the model is indeed capable of predicting the CME arrival time with high accuracy, even higher than the original J. Schmidt and Cairns (2019) model.

4 Conclusions

In this paper we have further improved the work of J. Schmidt and Cairns (2019), who achieved an excellent $0.9 \pm 1.9\,\mathrm{h}$ accuracy predicting the CME arrival times. Using the standard Richardson extrapolation technique, we have further improved the accuracy of the model to an average error to 0.2 ± 0.26 hours. We showed that the predictions are in the range 0.2 ± 0.4 hours with 95% confidence, and it is practically impossible that the good agreement between observations and simulation results obtained by J. Schmidt and Cairns (2019) was simply a lucky coincidence.

Given the low computational cost of the model and the fact that it relies on readily available real time observations, we believe that this break-through result can improve the current CME arrival prediction accuracy by more than an order of magnitude, and provide reliable and timely forecast for the space weather affected infrastructure operators, as well as enthusiasts of aurora observations.

5 Open Research

All data used in this paper are contained in Table 1. The Space Weather Modeling Framework including (BATS-R-US/AWSoM) is an open-source code available at https://github.com/MSTEM-QUDA with a full version history.

References

Amerstorfer, T., Hinterreiter, J., Reiss, M. A., Möstl, C., Davies, J. A., Bailey, R. L., … Harrison, R. A. (2021). Evaluation of CME arrival prediction using ensemble modeling based on heliospheric imaging observations. *Space Weather*, *19*, e2020SW002553. doi: 10.1029/2020SW002553

Arge, C., & Pizzo, V. (2000). Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates. *J. Geophys. Res.*, *105*, 10465-10479. doi: 10.1029/1999JA000262

Cohen, O., Sokolov, I., Roussev, I., Arge, C., Manchester, W., Gombosi, T., …

Velli, M. (2007). A semi-empirical magnetohydrodynamical model of the solar wind. *Astrophys. J.*, *654*, L163-L166. doi: 10.1086/511154

Odstrčil, D., & Pizzo, V. J. (1999a). Three-dimensional propagation of CMEs in a structured solar wind flow, 1, CME launched within the streamer belt. *J. Geophys. Res.*, *104*, 483–492. doi: 10.1029/1998JA900019

Odstrčil, D., & Pizzo, V. J. (1999b). Three-dimensional propagation of coronal mass ejections in a structured solar wind flow, 2, CME launched adjacent to the streamer belt. *J. Geophys. Res.*, *104*, 493–503. doi: 10.1029/1998JA900038

Powell, K., Roe, P., Linde, T., Gombosi, T., & De Zeeuw, D. L. (1999). A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *J. Comput. Phys.*, *154*, 284-309. doi: 10.1006/jcph.1999.6299

Richardson, L. F., & Gaunt, J. A. (1927). The deferred approach to the limit. *Philos. Trans. of the Royal Soc.*, *226*, 299. doi: 10.1098/rsta.1927.0008

Riley, P., Mays, M. L., Andries, J., Amerstorfer, T., Biesecker, D., Delouille, V., . . . Zhao, X. (2018). Forecasting the arrival time of coronal mass ejections: Analysis of the ccmc cme scoreboard. *Space Weather*, *16*(9), 1245-1260. doi: 10.1029/2018SW001962

Schmidt, J., & Cairns, I. (2019). Hit or miss, arrival time, and b_z orientation predictions of bats-r-us cme simulations at 1 au. *arXiv*, *1905*, e08961. doi: 10.48550/arXiv.1905.08961

Schmidt, J. M., & Cairns, I. H. (2014). Type ii solar radio bursts predicted by 3d mhd cme and kinetic radio emission simulations. *J. Geophys. Res.*, *119*, 69. doi: 10.1002/2013JA019349

Schmidt, J. M., & Cairns, I. H. (2016). Quantitative prediction of type ii solar radio emission from the sun to 1 au. *Geophys. Res. Lett.*, *43*, 50. doi: 10.1002/2015GL067271

Schmidt, J. M., Cairns, I. H., Cyr, O. C. S., Xie, H., & Gopalswamy, N. (2016). Cme flux rope and shock identifications and locations: Comparison of white light data, graduated cylindrical shell (gcs) model, and mhd simulations. *J. Geophys. Res.*, *121*. doi: 10.1002/2015JA021805

Schmidt, J. M., Cairns, I. H., & Hillan, D. S. (2013). *Astrophys. J.*, *773*, L30. doi: 10.1088/2041-8205/773/2/L30

Tóth, G., van der Holst, B., Sokolov, I. V., Zeeuw, D. L. D., Gombosi, T. I., Fang, F., ... Opher, M. (2012). Adaptive numerical algorithms in space weather modeling. *J. Comput. Phys.*, *231*, 870–903. doi: 10.1016/j.jcp.2011.02.006

van der Holst, B., Sokolov, I., Meng, X., Jin, M., Manchester, W. B., Tóth, G., & Gombosi, T. I. (2014). Alfvén wave solar model (AWSOM): Coronal heating. *Astrophys. J.*, *782*, 81. doi: 10.1088/0004-637X/782/2/81

Wold, A. M., Mays, M. L., Taktakishvili, A., Jian, L. K., Odstrcil, D., & MacNeice, P. (2018). Verification of real-time wsa-enlil+cone simulations of cme arrival-time at the ccmc from 2010 to 2016. *J. Space Weather Space Clim.*, *8*, A17. doi: 10.1051/swsc/2018005

Predicting Coronal Mass Ejection Arrival Times with Thirty-Minute Accuracy

Gábor Tóth

Dept. of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA

Abstract

J. Schmidt and Cairns (2019) have recently shown that they can predict Coronal Mass Ejection (CME) arrival times with an accuracy of 0.9±1.9 hours for four separate events. They also showed that the accuracy gets better with increased grid resolution. Here, we further improve these results by using the Richardson extrapolation (Richardson & Gaunt, 1927), which is a standard technique in computational fluid dynamics, and predict the CME arrival time with 0.2 ± 0.26 hours accuracy. The CME arrival time errors of the new model lie in a 95% confidence interval $[-0.21, 0.61]$ h. We also show that the probability of getting these accurate arrival time predictions with a model with a standard deviation exceeding 2 hours is less than 0.1%, indicating that the excellent results cannot be due to random chance, and the Richardson extrapolation has indeed improved the original model by J. Schmidt and Cairns (2019). This unprecedented accuracy is about 40 times better than the current state-of-the-art prediction of CME arrival times with an average error of about ±10 hours. The new model uses information available within a few hours after the CME eruption and it can run much faster than real-time on a couple of CPU cores. Based on the result, we recommend the new model to be transitioned to operations as soon as possible to better protect our space-born and ground-based assets from the harmful effects of space weather.

1 Introduction

Predicting the propagation of Coronal Mass Ejections (CMEs) and their arrival time at Earth has been a major goal of space weather prediction for decades. The ENLIL model (Odstrčil & Pizzo, 1999a, 1999b), for example, solves the ideal magnetohydrodynamic (MHD) equations from about $20\,R_s$ (solar radii) to the Earth orbit and beyond. The inner boundary conditions are provided by the Wang-Sheeley-Arge (WSA) model (Arge & Pizzo, 2000). CMEs are initiated with the empirical CONE model based on flare observations and coronal white light images. Another approach is followed by the Alfvén Wave Solar atmosphere Model (AWSoM) (van der Holst et al., 2014) that is based on the BATS-R-US MHD code (Powell et al., 1999; Tóth et al., 2012), also widely used to model the solar corona, the heliosphere and the eruption and propagation of CMEs from the surface of the Sun (initiated by a flux rope model) to Earth and beyond. AWSoM solves the MHD equations extended with solar wind heating and acceleration due to Alfvén wave turbulence, radiative cooling and heat conduction. However, these first-principles models can only achieve about 10-hour accuracy predicting the CME arrival time (Wold et al., 2018, cf.). More recently, empirical and neural network based models were applied to this problem, but the typical error remains about ± 10 hours (Riley et al., 2018; Amerstorfer et al., 2021, cf.).

J. Schmidt and Cairns (2019) have used an earlier version of BATS-R-US developed by Cohen et al. (2007), which relies on a spatially varying polytropic index derived from the Wang-Sheely-Arge (WSA) model (Arge & Pizzo, 2000) and achieved an unprecedented accuracy for predicting the CME arrival time: 0.9 ± 1.9 hours. They describe their procedure of setting up the CME simulations using only information that is available prior to and within a few hours after the CME eruptions: the WSO magnetogram, the CME speed estimated from the CME Analysis Tool (CAT) using STEREO/LASCO C3 coronagraph images, and prior L1 observations used for the WSA model and in turn for BATS-R-US. In addition, we have learned from the authors that the simulations were performed on a couple of CPU cores and they managed to run the model about three times faster than real time. This means that their procedure is undoubtedly suitable for operational use.

Unfortunately, J. Schmidt and Cairns (2019) have only published their work in form of a preprint on arxiv, so it did not get the attention that it deserves. In fact, their work has been peer-reviewed and accepted for publication in one of the leading journals, and the actual publication was only delayed for some minor technical reasons. As it is explained by J. Schmidt and Cairns (2019), the "setup and analysis is refined from our earlier work simulating type II radio bursts and CMEs", which in fact resulted in four peer-reviewed and published works (J. M. Schmidt et al., 2013; J. M. Schmidt & Cairns, 2014, 2016; J. M. Schmidt et al., 2016). Therefore the research contained in (J. Schmidt & Cairns, 2019) can be safely considered to have the same quality and rigor as these prior publications. It is therefore imperative to make sure that this research is not forgotten and benefit the entire space weather community.

Looking at Figure 4 in (J. Schmidt & Cairns, 2019), reproduced here as Figure 1, we have noticed that the distances between the observations (diamonds) and the model predictions obtained on two different computational grids (squares and stars) form a distinctive pattern: the distances between the three symbols appear to be approximately

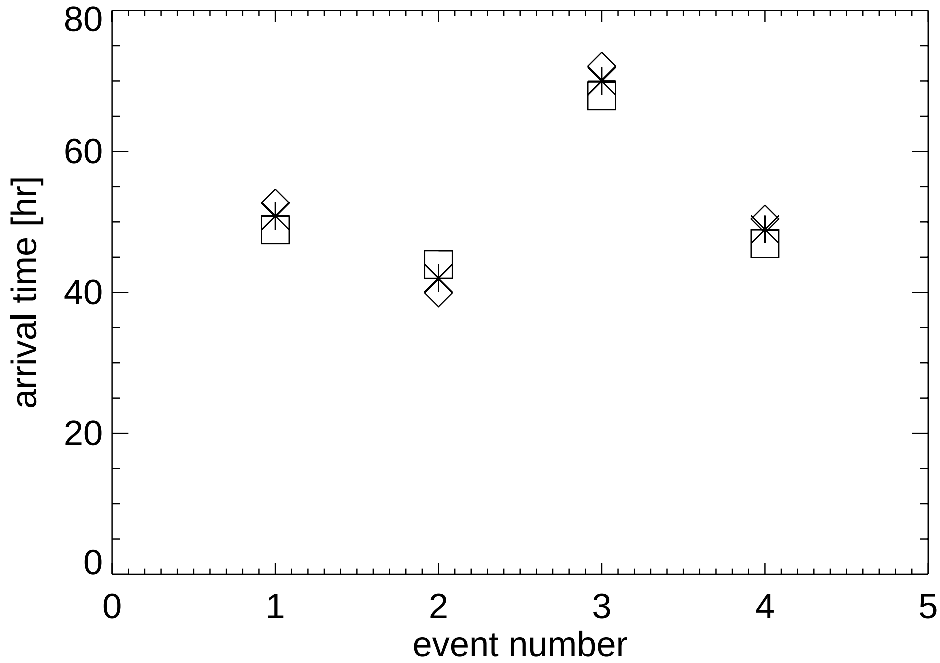

Figure 1. Observed and predicted arrival times at 1 au of four CME events (4 Sep 2017, 6 Sep 2017, 12 Feb 2018, and 29 Nov 2013 CME) recreated from (J. Schmidt & Cairns, 2019). The diamonds show observed arrival times, the squares and stars are simulation results at level 2 and level 5 grid refinements, respectively.

the same for all four events displayed. The main idea in this paper is that this fact can be exploited to obtain an even more accurate estimate of the CME arrival time.

The new arrival time estimate based on the Richardson extrapolation (Richardson & Gaunt, 1927) has a bias and standard deviation of 0.2 ± 0.26 hours, which is significantly better than the 0.9 ± 1.9 hours obtained by J. Schmidt and Cairns (2019). We will also show that the agreement between observations and simulations cannot be attributed to luck. Since the four events happened in different years and/or have very different arrival times covering a wide range from about 40 hours to 72 hours, the technique must be applicable to most CMEs. This means that the model should provide extremely reliable and accurate information for operational space weather forecasters, which is important for our national security and human safety.

2 Predicting CME arrival times

To perform a quantitative evaluation of the results presented in Figure 4 of (J. Schmidt & Cairns, 2019), we have digitized the figure and put the observed and simulated arrival times (relative to the eruption time) into Table 1. These values were also used to produce Figure 1 confirming that the values were extracted correctly.

Table 1. Simulated and observed CME arrival times for four events from Figure 4 in (J. Schmidt & Cairns, 2019). The times are measured in hours from the eruption time. The error is the absolute value of the difference between the observed and simulated times.

ID	Date	Observed	Model1	Model2	Error1	Error2	Error1/Error2
1	Sep 04, 2017	52.676	48.873	50.845	3.803	1.831	2.077
2	Sep 06, 2017	39.929	43.943	42.000	4.014	2.071	1.938
3	Feb 12, 2018	72.112	67.887	69.979	4.225	2.133	1.981
4	Nov 29, 2013	50.422	46.901	48.943	3.521	1.479	2.381
Average					3.891	1.8735	2.094

The errors Error1 and Error2 of the two models Model1 and Model2, corresponding to Refinement Level 2 and 5 in (J. Schmidt & Cairns, 2019), are remarkably constant across the four events, and the ratio of the errors is approximately 2.1. Using the idea of the Richardson extrapolation, which improves the numerical accuracy by estimating the exact solution from numerical solutions at two different grid resolutions, we construct the following formula for the extrapolated arrival time:

$$T_R = 2T_2 - T_1 \tag{1}$$

where T_1 and T_2 are the arrival times predicted by models 1 and 2. This provides the smallest error compared to observations and it is also consistent with the Richardson extrapolation for a first order accurate scheme. Indeed, the numerical scheme for discontinuities, like the CME shock, is only first order accurate.

3 Statistical Analysis and Probability Estimates

Table 2 shows that the mean absolute error of the extrapolated arrival time is about 0.218 hours, which is useful information, but not suitable for statistical analysis. To bet-

Table 2. Observed and extrapolated CME arrival times for four events. The times are measured in hours from the eruption time. The last column is the absolute value of the error.

ID i	Date	Observed T_i	Extrapolated $T_{i,R}$	Error $T_{i,R} - T_i$
1	Sep 04, 2017	52.676	52.817	0.141
2	Sep 06, 2017	39.929	40.057	0.128
3	Feb 12, 2018	72.112	72.071	-0.041
4	Nov 29, 2013	50.422	50.985	0.562
Mean absolute error				0.218
Mean $\pm$ one standard deviation				0.198 ± 0.257

ter quantify the performance of the new model, we calculate an unbiased estimate and a 95% confidence interval for the arrival time errors.

The sample size is $N = 4$. The average of the errors, the bias, is

$$B = \frac{1}{N}\sum_{i=1}^{N}(T_{i,R} - T_i) = 0.198\,h \tag{2}$$

and the sample variance S is

$$S = \sqrt{\frac{\sum_{i=1}^{N}(T_{i,R} - T_i - B)^2}{N-1}} = 0.257\,\mathrm{h} \tag{3}$$

where T_i is the observed arrival time for event i and $T_{i,R}$ is the Richardson extrapolated time calculated from Equation 1. The 95% confidence interval for the error $T_R - T$ is $B \pm tS/\sqrt{N}$, where $t = 3.182$ from the T-distribution for $p = 0.025$ and $N - 1 = 3$ degrees of freedom:

$$(T_R - T) \in [-0.211, 0.607]\,\mathrm{h} \tag{4}$$

We conclude that there is a 95% chance that the model will produce arrival time predictions with errors less than 37 minutes, while the average error is only 12 minutes.

Finally, it is important to check if the small errors in Table 2 are statistically significant, or they can be attributed to simple luck. Let us assume that the new model with the extrapolation has no bias, $\mu = 0$, and its standard deviation is $\sigma = 2\,\mathrm{h}$. The quantity

$$X^2 = \frac{\sum_{i=1}^{N}(T_{i,R} - T_i)^2}{\sigma^2} = 0.0887 \tag{5}$$

follows the $\chi^2(N,p)$ distribution since the mean value is assumed to be known. For $N = 4$, we find that there is only $p = 0.1\%$ chance that $X^2 \leq 0.0887$ by pure luck. If σ was larger than 2 hours, this probability would be even less. We can safely conclude that the model is indeed capable of predicting the CME arrival time with high accuracy, even higher than the original J. Schmidt and Cairns (2019) model.

4 Conclusions

In this paper we have further improved the work of J. Schmidt and Cairns (2019), who achieved an excellent 0.9 ± 1.9 h accuracy predicting the CME arrival times. Using the standard Richardson extrapolation technique, we have further improved the accuracy of the model to an average error to 0.2 ± 0.26 hours. We showed that the predictions are in the range 0.2 ± 0.4 hours with 95% confidence, and it is practically impossible that the good agreement between observations and simulation results obtained by J. Schmidt and Cairns (2019) was simply a lucky coincidence.

Given the low computational cost of the model and the fact that it relies on readily available real time observations, we believe that this break-through result can improve the current CME arrival prediction accuracy by more than an order of magnitude, and provide reliable and timely forecast for the space weather affected infrastructure operators, as well as enthusiasts of aurora observations.

5 Open Research

All data used in this paper are contained in Table 1. The Space Weather Modeling Framework including (BATS-R-US/AWSoM) is an open-source code available at https://github.com/MSTEM-QUDA with a full version history.

References

Amerstorfer, T., Hinterreiter, J., Reiss, M. A., Möstl, C., Davies, J. A., Bailey, R. L., ... Harrison, R. A. (2021). Evaluation of CME arrival prediction using ensemble modeling based on heliospheric imaging observations. *Space Weather*, *19*, e2020SW002553. doi: 10.1029/2020SW002553

Arge, C., & Pizzo, V. (2000). Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates. *J. Geophys. Res.*, *105*, 10465-10479. doi: 10.1029/1999JA000262

Cohen, O., Sokolov, I., Roussev, I., Arge, C., Manchester, W., Gombosi, T., ...

Velli, M. (2007). A semi-empirical magnetohydrodynamical model of the solar wind. *Astrophys. J.*, *654*, L163-L166. doi: 10.1086/511154

Odstrčil, D., & Pizzo, V. J. (1999a). Three-dimensional propagation of CMEs in a structured solar wind flow, 1, CME launched within the streamer belt. *J. Geophys. Res.*, *104*, 483–492. doi: 10.1029/1998JA900019

Odstrčil, D., & Pizzo, V. J. (1999b). Three-dimensional propagation of coronal mass ejections in a structured solar wind flow, 2, CME launched adjacent to the streamer belt. *J. Geophys. Res.*, *104*, 493–503. doi: 10.1029/1998JA900038

Powell, K., Roe, P., Linde, T., Gombosi, T., & De Zeeuw, D. L. (1999). A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *J. Comput. Phys.*, *154*, 284-309. doi: 10.1006/jcph.1999.6299

Richardson, L. F., & Gaunt, J. A. (1927). The deferred approach to the limit. *Philos. Trans. of the Royal Soc.*, *226*, 299. doi: 10.1098/rsta.1927.0008

Riley, P., Mays, M. L., Andries, J., Amerstorfer, T., Biesecker, D., Delouille, V., . . . Zhao, X. (2018). Forecasting the arrival time of coronal mass ejections: Analysis of the ccmc cme scoreboard. *Space Weather*, *16*(9), 1245-1260. doi: 10.1029/2018SW001962

Schmidt, J., & Cairns, I. (2019). Hit or miss, arrival time, and b_z orientation predictions of bats-r-us cme simulations at 1 au. *arXiv*, *1905*, e08961. doi: 10.48550/arXiv.1905.08961

Schmidt, J. M., & Cairns, I. H. (2014). Type ii solar radio bursts predicted by 3d mhd cme and kinetic radio emission simulations. *J. Geophys. Res.*, *119*, 69. doi: 10.1002/2013JA019349

Schmidt, J. M., & Cairns, I. H. (2016). Quantitative prediction of type ii solar radio emission from the sun to 1 au. *Geophys. Res. Lett.*, *43*, 50. doi: 10.1002/2015GL067271

Schmidt, J. M., Cairns, I. H., Cyr, O. C. S., Xie, H., & Gopalswamy, N. (2016). Cme flux rope and shock identifications and locations: Comparison of white light data, graduated cylindrical shell (gcs) model, and mhd simulations. *J. Geophys. Res.*, *121*. doi: 10.1002/2015JA021805

Schmidt, J. M., Cairns, I. H., & Hillan, D. S. (2013). *Astrophys. J.*, *773*, L30. doi: 10.1088/2041-8205/773/2/L30

Tóth, G., van der Holst, B., Sokolov, I. V., Zeeuw, D. L. D., Gombosi, T. I., Fang, F., … Opher, M. (2012). Adaptive numerical algorithms in space weather modeling. *J. Comput. Phys.*, *231*, 870–903. doi: 10.1016/j.jcp.2011.02.006

van der Holst, B., Sokolov, I., Meng, X., Jin, M., Manchester, W. B., Tóth, G., & Gombosi, T. I. (2014). Alfvén wave solar model (AWSOM): Coronal heating. *Astrophys. J.*, *782*, 81. doi: 10.1088/0004-637X/782/2/81

Wold, A. M., Mays, M. L., Taktakishvili, A., Jian, L. K., Odstrcil, D., & MacNeice, P. (2018). Verification of real-time wsa-enlil+cone simulations of cme arrival-time at the ccmc from 2010 to 2016. *J. Space Weather Space Clim.*, *8*, A17. doi: 10.1051/swsc/2018005